Technical Approach to Developing a Spatial Crime Analysis System with ArcView GIS

David Nulph, INDUS Corporation
Jeff Burka, INDUS Corporation
Alex Mudd, U.S. Department of Justice

April 15, 1997

Abstract

Objective: The purpose of this paper is to discuss the methods used to develop and implement an ArcView based spatial crime analysis system which will allow police departments to geographically analyze incident-based data.

Background: The Criminal Division of the U.S. Department of Justice initiated two pilot projects to implement GIS for use in crime analysis at local police departments: Warrenton, VA and Montgomery County, MD. The crime analysis application was first installed at Warrenton, a very small police department, then with some enhancements and added capabilities at the Montgomery County Police Department (MCPD).

Application Requirements: The key requirements of the system include an easy to use database query interface, the ability to produce several varieties of analytical maps and reports, and the flexibility to be installed on different systems with a minimum of re-programming.

Technical Approach: The core application functionality was built around Environmental Systems Research Institute's ArcView 3.0 using Avenue. Additional functionality was provided by the ArcView Spatial Analyst extension, which allowed the rapid inclusion of features such as hot-spot analysis and crime density surface map generation.

To achieve all the application requirements, several additional components were integrated into the SCAS ArcView environment: Visual Basic forms are used to handle user interface needs and communicate to ArcView via Dynamic Data Exchange (DDE). Report generation capability is achieved by directly linking ArcView to the Crystal Reports report generator engine. Database connectivity issues are resolved by standardizing the database query mechanism on ODBC (Open DataBase Connectivity) SQL (Structured Query Language).

 The resulting application is currently in use at the two police departments, and the Department of Justice, Criminal Division GIS staff continues to enhance the application to optimize it to the police departments' needs.

Background

In 1994, the U.S. Department of Justice, Criminal Division initiated two pilot projects to implement GIS spatial crime analysis systems for use at the local law enforcement level. The DOJ chose two sites for these projects, a very small police department in the Town of Warrenton, Virginia, and the much larger police department in Montgomery County, Maryland. With programming and application development support from INDUS Corporation, these two GIS crime analysis implementations are underway, with additional sites under review. This paper will discuss the technical approach to developing the spatial crime analysis application for the DOJ Crime Analysis Systems Project.

Application Requirements

The requirements of the Spatial Crime Analysis System (SCAS) are as follows:

- Run on a desktop PC using the Windows 95 operating system.

- Use existing police department incident and arrest data sources.

- Recognizing the Department of Justice intent to install the application in multiple police departments, the application must be as portable as possible.

- Be easy to use. In all cases, the users would be largely unfamiliar with GIS software. The application would need to handle the more complicated aspects of GIS analysis, and 'guide' the user through the analysis. The goal was to develop an application that permits crime analysts to produce finished, ready to plot maps, charts and reports, yet shield them from the complexities of GIS database analysis.

- Provide a core foundation of database query and mapping tools which could be enhanced as the need arose.

Software Components

The core functionality of SCAS was designed around Environmental System Research Institute's ArcView 3.0 product. Avenue, the ArcView development language, was used to customize and automate the application. ArcView was chosen because it runs on a desktop PC, is customizable with Avenue, and is relatively easy to use. The ArcView Spatial Analyst extension was also used to facilitate the generation of 3D surfaces which are used for hot-spot analysis. Since ArcView 3.0 has no native form menu development capabilities, form menus developed with Visual Basic were integrated into the application via Dynamic Data Exchange (DDE). The Crystal Reports report generation engine, which is included with Visual Basic, was also used to add report generation capabilities.

Component Functionality

Open DataBase Connectivity (ODBC)

One of the key components of SCAS is utilization of ODBC. Almost all SCAS external database access was implemented using ODBC. ODBC provides a standard layer between the application and the data sources (Figure 1).

ODBC is an industry standard application programming interface which allows applications to communicate with various database systems and even to treat some non-database files, such as spreadsheets, as if they were database tables. ODBC is supported on numerous platforms, and supports numerous database vendors. The ODBC Application Programmers Interface (API) allows a programmer to communicate through the Structured Query Language (SQL) with any database for which there is an ODBC driver - effectively insulating the programmer from vendor-specific database peculiarities such as database file formats and query mechanics.

ArcView supports ODBC both at the standard interface level as well as via Avenue. Users may open a connection to an external database and then interact with that database in numerous ways, whether inserting new records into the database, updating existing records, or using a query to bring records from the database into an ArcView VTab. All interaction with the external database is done with queries written in SQL.

There are several advantages to using ODBC for virtually all queries of the incident database. The main advantage is simply that data does not need to be exported from the police department's database in order to be used with ArcView. The department maintains its existing database with entry, query, and reporting features, while SCAS is able to seamlessly access the same database to fulfill its data needs. As a matter of practicality, the geocoded point data is also stored as part of the database. ArcView is able to create a point shapefile directly from VTabs generated by ODBC queries of the database, using the XYName subclass of the SrcName class.

Theoretically, having the ODBC layer between SCAS and the database will permit the application to access data from a wide variety of DataBase Management Systems (DBMS) systems (currently DataFlex and MS Access, and Oracle in the future), with minimal modifications to the application itself, making the product more portable among police departments. The modifications required for a new site will primarily address differences in database structure such as field names and table relationships.
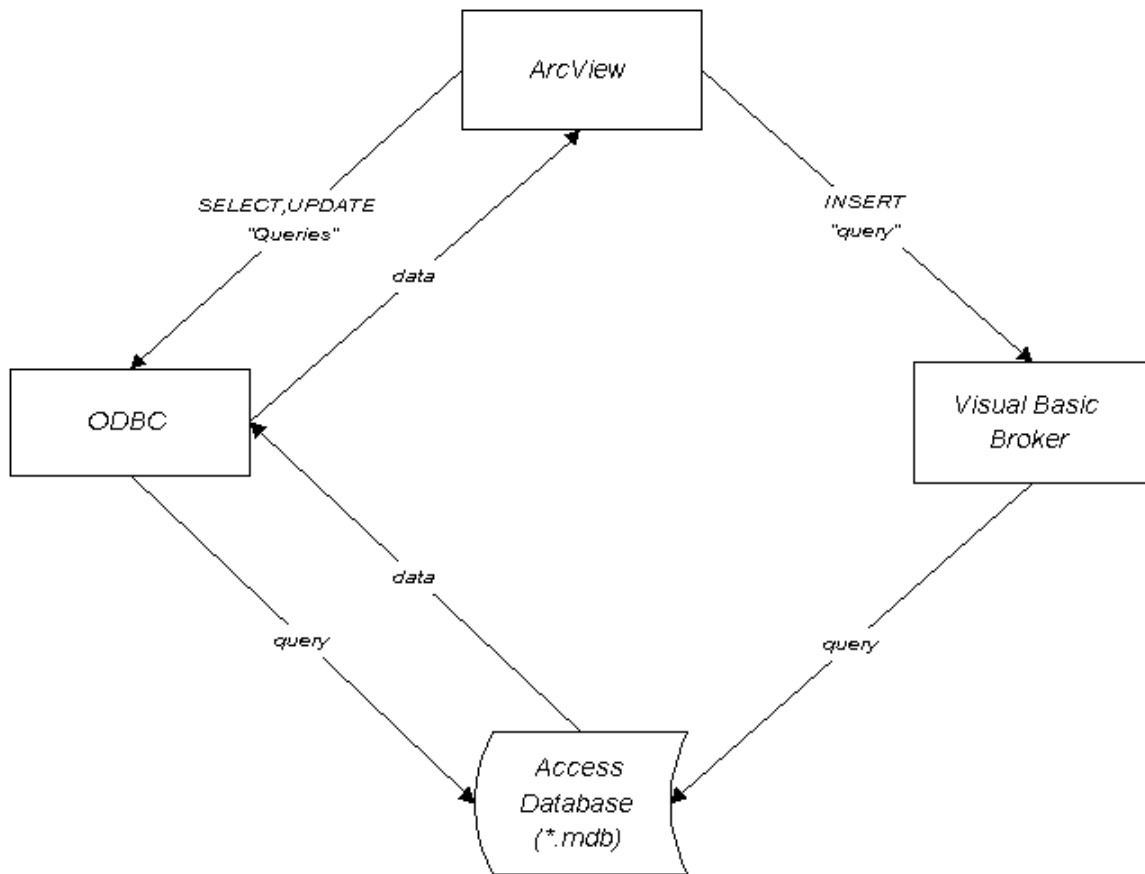
Figure 1: Data Flow Between ArcView and the Access-based Database

Visual Basic Form Menus / DDE

The core interface elements for many of the analysis tools were developed using Microsoft Visual Basic 4.0. A series of Visual Basic forms perform various user input and data gathering tasks. The forms are invoked by the ArcView application using a simple 'system.execute' request to the form executable file. Almost all of the forms are designed to subsequently pass information back to ArcView, and this is accomplished with a DDE link back to the ArcView session as a client. An argument is passed with the DDE request in the form of a pipe delimited text string. The receiving ArcView script then parses the argument string, and proceeds with the procedure at hand.

A very useful component used in several of the Visual Basic forms is the TreeView custom control provided with the Visual Basic distribution. This control was used extensively in developing the Primary Database Query form (Figure 2) . This control permits the construction of hierarchical pick lists, greatly simplifying tasks such as the selection of crime classification codes.

Figure 2: Primary Query Form

In addition, components from the Crescent QuickPak Professional, published by Progress Software, are utilized to add functionality such as calendars and enhanced combo boxes, as well as enhanced string manipulation and other programmer-level functions.

Visual Basic Daemons/Brokers: *sizedaemon* and *geocode*

Visual Basic was primarily used to develop the user interface forms and dialog boxes for SCAS, but Visual Basic was also used to create two programs which have nothing to do with the user interface. Both of these programs address functions not available in the core ArcView API, but nonetheless are required to perform important functions for SCAS to operate properly.

The first of these is the "sizedaemon." SCAS uses a great deal of the screen, even on a 1024x768 screen size. When running in a lower resolution, it is desirable to move and resize a number of aspects of the project. While it is possible to use the application object to determine the ArcView window size, there is no way to actually trap the change in window size and thereby run an update script. While an update script might have been attached to an object capable of storing such a property, there was no way to guarantee that the update script would be run each time the user changed the ArcView main window size.

The size daemon is a very small Visual Basic program which sits quietly in the background. Every few seconds, the daemon uses a Windows API call to poll the main ArcView application window size to see if it has changed. If so, the size daemon notifies ArcView that its size has changed via

DDE, telling ArcView to run a script which repositions a number of items in a few of the DocGUIs and resizes certain windows as necessary. Virtually no performance degradation has been noticed while running the size daemon. It very quickly goes back to sleep if the ArcView window size has not been changed; no spikes are visible in the Windows 95 CPU monitor when the size daemon is running.

The second Visual Basic program that is not used for user interface development is a small broker which sits between ArcView and ODBC and handles those transactions in which records are inserted into the database ([Figure 1](#)). When the system was first tested in a multi-user situation, with the master database stored on a file server, there were tremendous slow-downs, large quantities of modal warning messages, and even crashes occurring when two users were engaged in geocoding records.

Although ArcView's ODBC capabilities are fairly good, it was discovered that there is no way to trap errors returned from ODBC queries. In effect, ODBC was generating errors, ArcView was noting the error, and then warning the user - occasionally by displaying up to 1,000 modal dialog boxes in a row. With no way to trap these errors and deal with them via Avenue, a Visual Basic program was written to accept the records to be inserted from ArcView and pass them along to the database, trapping errors, repeating attempts, and generally speeding things up for concurrent users.

The Visual Basic broker does slow performance for a single user inserting geocodes into the database; in a 1200 record sample, the slow-down was approximately 12%. But the improvement in multi-user access has far outweighed the speed degradation caused by the broker.

Crystal Reports Report Generation

Another component of the Visual Basic distribution integrated into SCAS was the Crystal Reports report generation engine. This was necessary because SCAS required fairly sophisticated reports to accompany the maps. To accomplish this, the Crystal Reports Print Engine (CRPE32.DLL) is initialized from within Avenue. This makes available to ArcView the entire suite of Crystal Reports functions, which can then be called directly.

 To generate a report, a previously prepared report form is initialized and sent a .dbf file exported by an Avenue Script. The report can then be previewed and sent to a printer. As requested by the crime analysts, the report generation component of SCAS was designed to be as generic as possible, and permit the generation of reports independently of maps. This independence permits SCAS to be used solely for report generation purposes, as there are many occasions when an analyst needs only reports.

The major drawback of this system is that a previously designed report form must be available. This restricts the tables which can be printed, and forces a report form redesign if the database table structure changes, particularly fieldnames.

Microsoft Help System

The Microsoft help compiler was used to develop an on-line, indexed, keyword searchable help system for SCAS. Several help system Microsoft Word macros were tested, and found to be inadequate for various reasons. The Microsoft help compiler, though somewhat cumbersome, permitted all of the help functionality to be integrated relatively easily. The help file is called from within the application environment with the command 'system.execute (Winhelp "scashelp.hlp","").' A context-sensitive help link was also provided for the SCAS buttons and tools. This feature was implemented via the 'Help Topic' property of the buttons and tools. By selecting

the context help button and pointing at a button or tool, the user is automatically taken to the correct topic in the help system.

Sample Application Functionality

Pin-Maps and Summaries

The basis for most of the analytical functionality, and the default 'first step' in a crime analysis operation, is the generation of pin-maps. These maps, an analogue to the paper maps and pushpins that police departments have used (and still use) for simple geographic analysis provide a very simple window into the distribution of incidents in an area. The production of a pin-map with SCAS is simple: an analyst phrases a query with the Primary Query form (Figure 2), and submits the form with the 'Submit Query' button. SCAS performs the query, returns a set of records, geocodes any previously ungeocoded records and places the resulting points on the map. The resulting map is relatively easy to interpret, and is a good starting point for subsequent analysis. A classification tool permits uniquely symbolizing each point by its crime classification code (Figure 6)

As part of the pin-map generation process, a choropleth summary can automatically be produced. This option counts the incident points in each polygon (beat, district, ZIP, etc.) and automatically classifies and symbolizes the polygons accordingly. The current version of SCAS requires the user to perform normalized classifications by area or population manually, but such functionality will be automated in future versions of SCAS.

Geocoding (Manual and Automatic)

One of the most important functions of the system is geocoding. Without an accurately geocoded incident database, spatial analysis would be impossible. SCAS was developed so that most geocoding occurs automatically and is transparent to the user. Every time an analyst performs a query, the resulting recordset is scanned, previously un-geocoded records are extracted, and geocoded. The newly geocoded x, y coordinates are then stored back in the master database. The next time a query is performed, previously geocoded records will bypass the geocoding procedure, making the query-geocode process much more efficient.

Unfortunately, due to several factors such as data entry errors and an out of date street file, the geocode rate is typically around 90%. To improve the geocode rate, a mechanism was developed to permit the MCPD analysts to manually geocode unmatched records. The ESRI script MatchCoordinate was modified to eliminate duplicate addresses, and to commit new x, y records back to a master geocode table, stored in the master Access database. This tool permits the user to pick an address from a pop-up pick list, then point to a location on the map to assign a geocode.

Change Maps

Another fundamental need of the crime analysts is to be able to note trends in crime. In which areas is crime increasing? decreasing? staying the same? A 'change map' tool was created which assists analysts in performing trend analysis. The tool works by building two incident count maps by polygon, for selected time periods. The tool is designed to permit the comparison of any two independently selected periods of time. The second (later time period) map is then subtracted

from the first (earlier time period) map to produce a third map which shows the increase or decrease in incidents per polygon between the two time periods.

Surface-Derived Hot-Spots

One of the popular topics among crime analysts is hot-spot analysis. Hot-spots are areas of high concentrations of incidents within a police jurisdiction. There are many ways to define and analyze hot-spots for a given set of incident locations. The method used by SCAS is to incorporate ArcView Spatial Analyst and build a surface of incident density for a selected set of incident points, using the kernel function of Spatial Analyst's density surface generation routine. Once this surface is generated, it is relatively simple to reselect out the 'peaks' depicting hot-spots (Figure 3). A tool was also developed with Visual Basic to allow the analyst to adjust the cut-off point of the hot-spots via a sliding bar.
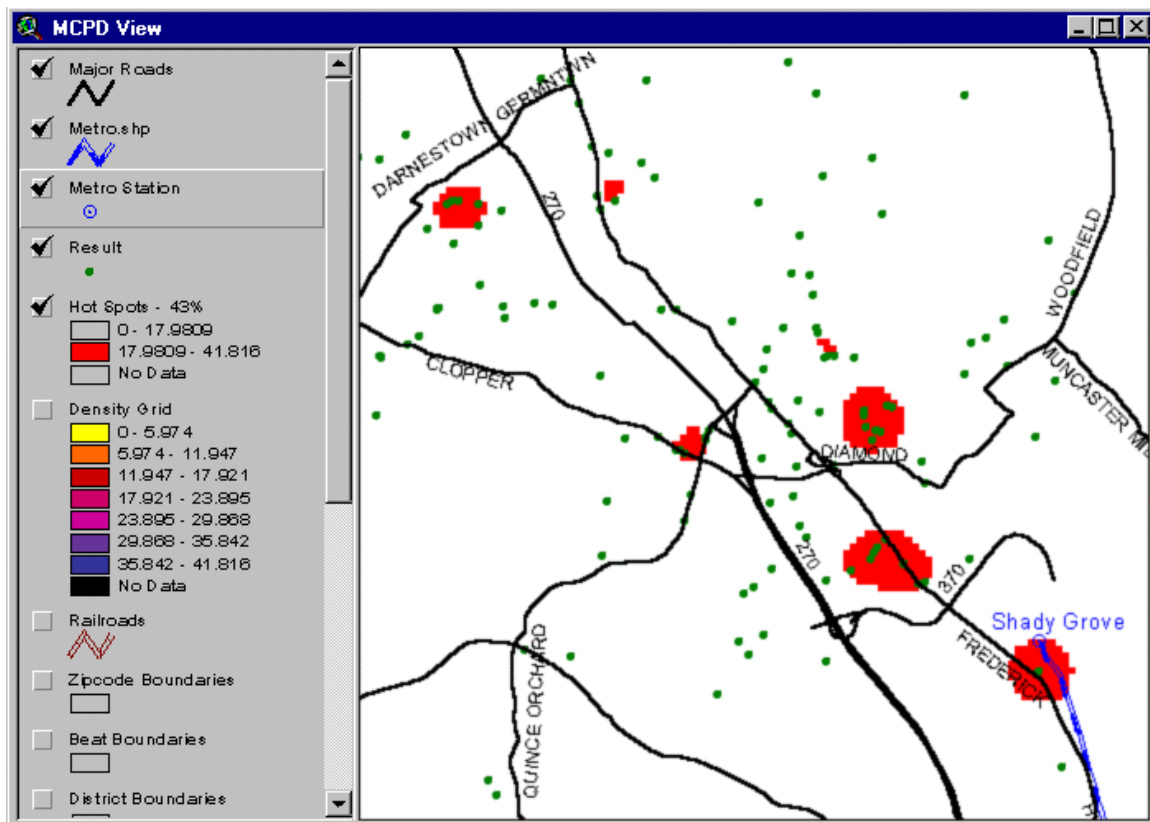

Figure 3: Hot Spots

Standard Deviation Ellipses

Another tool which can be used to assist in 'hot-spot' analysis is the standard deviation ellipse (SDE) generator. This tool generates a set of standard deviation ellipses based on the x, y distribution of a point theme, or a selected set of those points. SDEs have major and minor axes which match the standard deviations of the x and y coordinates of the points. The ellipses are also rotated to fit the slope of the distribution.

Regions based on standard deviation are often used in crime analysis in predictive models for serial crimes; if an analyst has determined that a series of crimes was likely committed by a single person or group, known crime locations can be used to help predict where and for what purpose law enforcement personnel can be deployed (Gottlieb, 1994). In situations where the power of computers and statistical software is not available, crime analysts resort to calculating geometric means for a distribution of crimes, and then using the X and Y standard deviations of that distribution to generate a series of rectangles around the points. For SCAS, an Avenue routine that generates standard deviation ellipses was written; this generic routine is then used by several different tools to generate one or more ellipses around distributions. These ellipses may then be displayed as graphics, or converted to a shapefile (Figure 4).
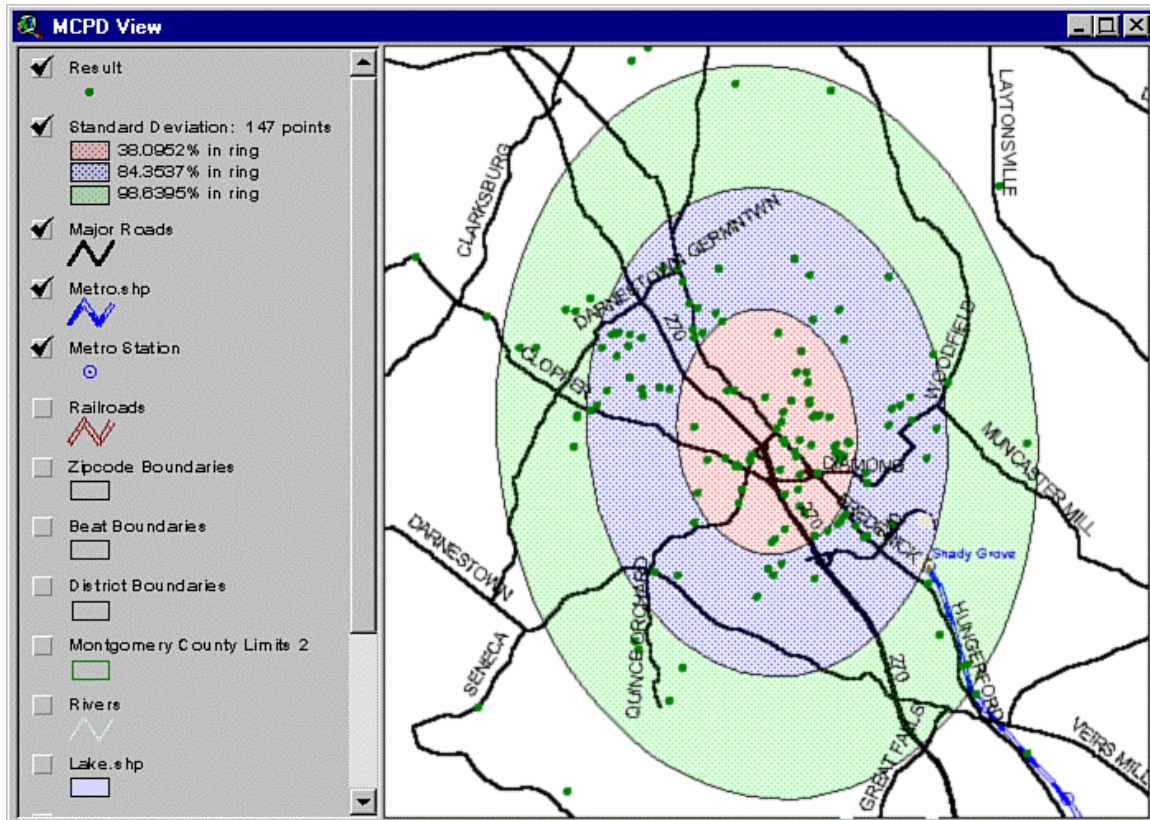

Figure 4: Standard Deviation Ellipses

Linkage Tools: Suspect Location and Recovered Vehicle

In many cases, the data one examines in crime analysis is better considered not in terms of single point features, but rather as dyads. A vehicle is stolen in one location and recovered in another. The location of a criminal's home may have some bearing on where he commits a crime. By hiding the connections between various types of origins and destinations, some patterns of behavior can be obscured.

SCAS contains a set of tools developed to address specific crime analysis needs in discovering and analyzing links between incident locations and suspects or stolen vehicles. These tools can be divided into two categories. The first operates on a per-feature basis and is based on two-way links between tables containing crime incident information and suspect or recovered vehicle

information. For a given set of incidents, a new point theme containing suspects associated with the incidents is created and linked to the original incident theme. Selecting a point feature in either theme causes the corresponding point in the other theme to be highlighted as well.

The second type of tool more explicitly shows the links between origins and destinations, by drawing the dyad as a pair of points with a line between them. SCAS has several symbolization options for these lines, including arrowheads indicating the destination end, and distance labeling along the line.

Temporal and Spatial Trend Charts

Two charting tools were developed which allow the crime analysts to graph incident occurrences both spatially and temporally. Both of the tools operate on a selected set of incident database records, which were the result of a database query.

The temporal (Figure 5) chart graphically displays in a line chart the number of incidents that occurred in one hour increments for the entire 24 hours of a day. Any records which do not have a time recorded are counted and displayed on the chart as the "25th hour." Up to three higher-level crime types (such as robbery, burglary, or sex crimes) can be charted simultaneously. This chart provides a very easy to interpret view of the temporal patterns of crime occurrence during the day.
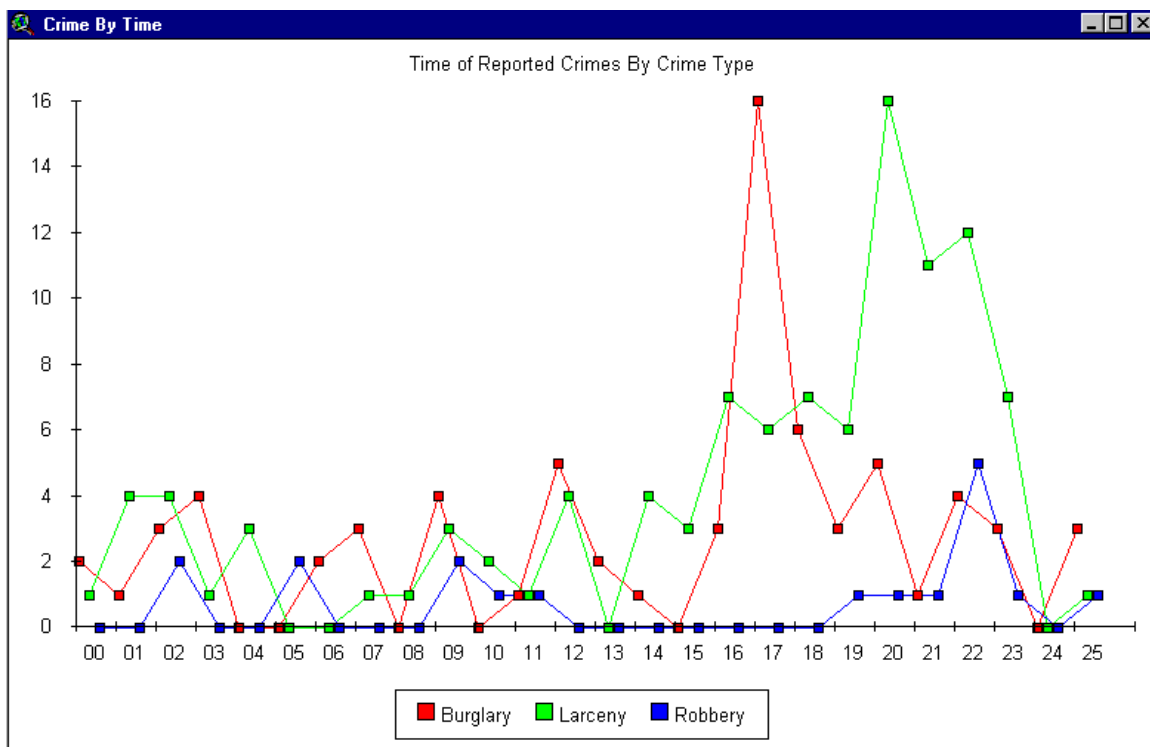

Figure 5: Automated Trend Charting

The spatial trend chart summarizes the table by counting the number of incidents for each Police Reporting Area (PRA) and/or police beat, then creates a bar chart. This chart is a very simple way to compare numbers of incidents of crime by polygon, and is a good supplement to maps which

have been created. In addition, a table for each chart type is created containing all of the data used to construct the charts.

Layout Generation

One of the primary responsibilities of any mapping package is, of course, to generate maps. Several scripts were written for SCAS which completely automate the generation of plotter-ready layouts. The goal of automating this process was to produce uniform maps, offer multiple page sizes (8.5"x11" to 34"x44") and page orientations (landscape and portrait). The layout generation tools create the layout based on the current main view window.

To accomplish these goals, several scripts were developed to ensure consistent layout generation. An update script was created that would resize the view document window if the view's length and width were not the same. This was necessary to ensure that changes made to the view's aspect ratio do not affect the proportions of the view in the layout.

Additional scripts were developed to create the layout templates for each page size and page orientation. If the user selects a template that does not exist within the project, SCAS will create the appropriate layout template. ESRI's multi-column legend script was utilized to create the multi-column legend in the portrait layouts.

To create a layout, the analyst first prepares the main View, then chooses 'Create Layout' from the 'Crime Tools' menu bar. Then, after answering a few dialog boxes for information such as the title and subtitles the finished layout is generated.
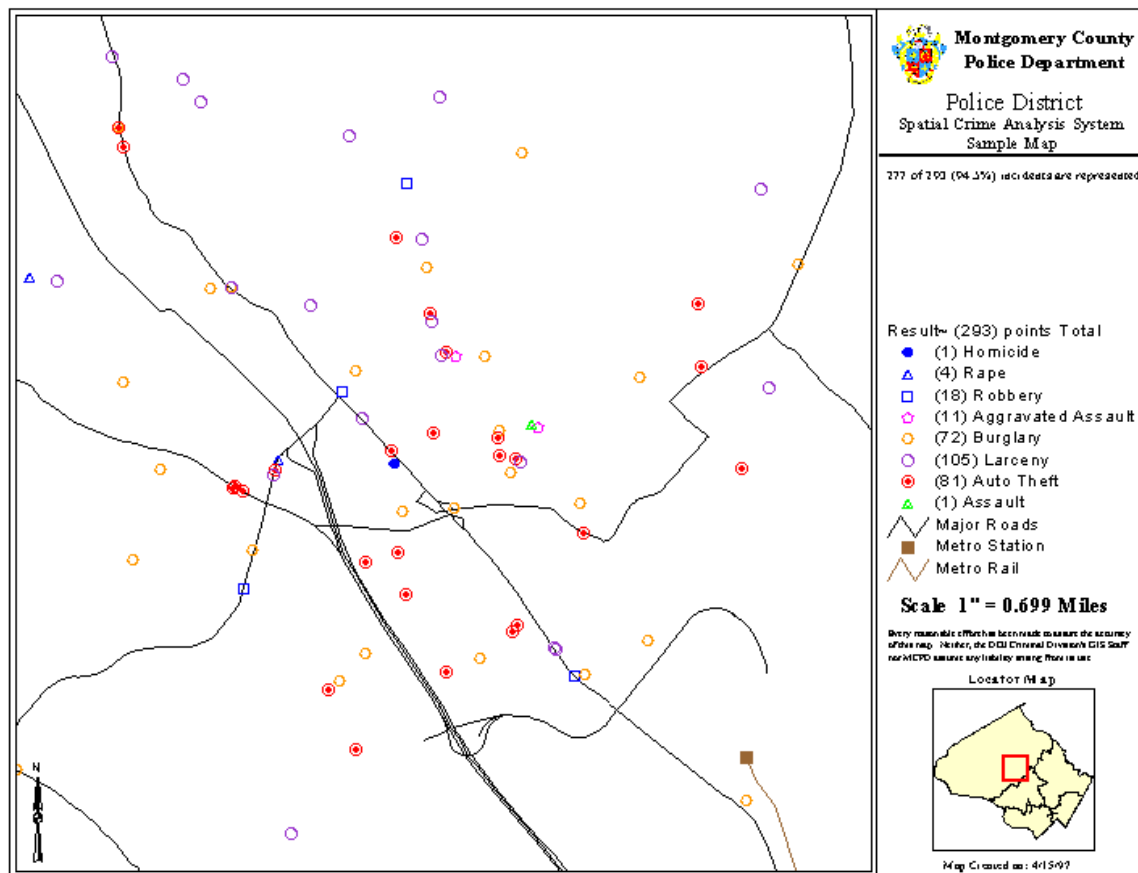
Figure 6: Automated Layout Generation

Future Plans

The current plans for future development of SCAS include continued automation and enhancement the application based on feedback from the currently operational sites.

Since the current implementation of SCAS was developed for use by crime analysts, one goal for future versions will be to develop tools which are appropriate for use at the 'officer on the street' level. The purpose of these tools will be to incorporate spatial crime analysis technology into the functions of police officers day-to-day work.

There are currently three additional Police Departments under review by the Department of Justice for SCAS implementation. As these sites are implemented, the knowledge gained at each site will be incorporated back into SCAS as a whole, the intent being to create a crime analysis application which can address the needs of as many police departments as possible.

Acknowledgments

Thanks to Robert Crump, Joe Pindell, and Vince Breneman for their contributions to the development of SCAS. Appreciation is also extended to the staffs of the Warrenton and Montgomery County Police Departments for their cooperation.

For Further Reading

The following references were used in the development of SCAS:

ArcView

Environmental Systems Research Institute (1996). *Avenue Customization and Application Development for ArcView.*

Razavi, Amir H., John Alexander and Valerie Warwick (1995). *ArcView/Avenue Programmer's Reference: Class Hierarchy Quick Reference and 101 Scripts.* Santa Fe: OnWord Press.

Crime Analysis

Block, Carolyn Rebecca and Margaret Dabdoub (Editors) (1993). *Workshop on Crime Analysis Through Computer Mapping Proceedings: 1993.* Chicago: Illinois Criminal Justice Information Authority.

Eck, John E. and David Weisburd (Editors) (1995). *Crime and Place, Crime Prevention Studies, Volume 4.* Monsey: Criminal Justice Press

Gottlieb, Steven, Sharon Arenberg, and Rag Singh (1994). *Crime Analysis from First Report to Final Arrest.* Alpha Publishing

Illinois Criminal Justice Information Authority (1988). *Spatial and Temporal Analysis of Crime Technical Manual.* Chicago: Illinois Criminal Justice Information Authority.

Taxman, Faye S. and J. Tom McEwen (1994). "High Tech Computer Mapping and Low-Tech Community Policing." *The Journal of Urban Technology*, Volume 2, Number 1, pages 85-103.

Visual Basic

Appleman, Daniel (1996). *Visual Basic Programmer's Guide to the Win32 API.* Emeryville: Ziff-Davis Press.

Bloom, Stuart and Don Kiely (1995)*. Visual Basic 4 Database How-To.* Corte Madera: Waite Group Press.

Environmental Systems Research Institute (1995). *ArcView Version 2 Working With Visual Basic.*

Jennings, Roger (1996). *Database Developers Guide with Visual Basic 4 Second Edition.* Indianapolis: Sams Publishing

Wang, Wallace (1995). *Visual Basic 4 For Windows For Dummies.* Foster City: IDG Books Worldwide, Inc.

Wang, Wallace (1996). *More Visual Basic 4 For Windows For Dummies.* Foster City: IDG Books Worldwide, Inc.

ODBC / SQL

Geiger, Kyle (1995). *Inside ODBC.* Redmond: Microsoft Press.

Groff, James R. and Paul N. Weinberg (1994). *LAN Times Guide to SQL.* Berkeley: Osborne

McGraw-Hill.;

MS Help Compiler

Microsoft Press (1995). *Microsoft Windows 95 Help Authoring Kit.* Redmond: Microsoft Press

Please direct inquiries regarding GIS at the Department of Justice, Criminal Division to:

John DeVoe
GIS Staff Chief
US Department of Justice, Criminal Division
1400 New York Avenue, Room 7120
Washington, DC 20005

Criminal.gis@usdoj.gov

Authors:

David Nulph
Senior GIS Analyst
INDUS Corporation
1953 Gallows Road, Suite 300
Vienna, VA 22182
(703)506-6700
dnulph@induscorp.com

Jeff Burka
GIS Analyst
INDUS Corporation
1953 Gallows Road, Suite 300
Vienna, VA 22182

jburka@induscorp.com

Alex Mudd
GIS Analyst
US Department of Justice, Criminal Division
1400 New York Avenue, Room 7120
Washington, DC 20005
criminal.gis@usdoj.gov